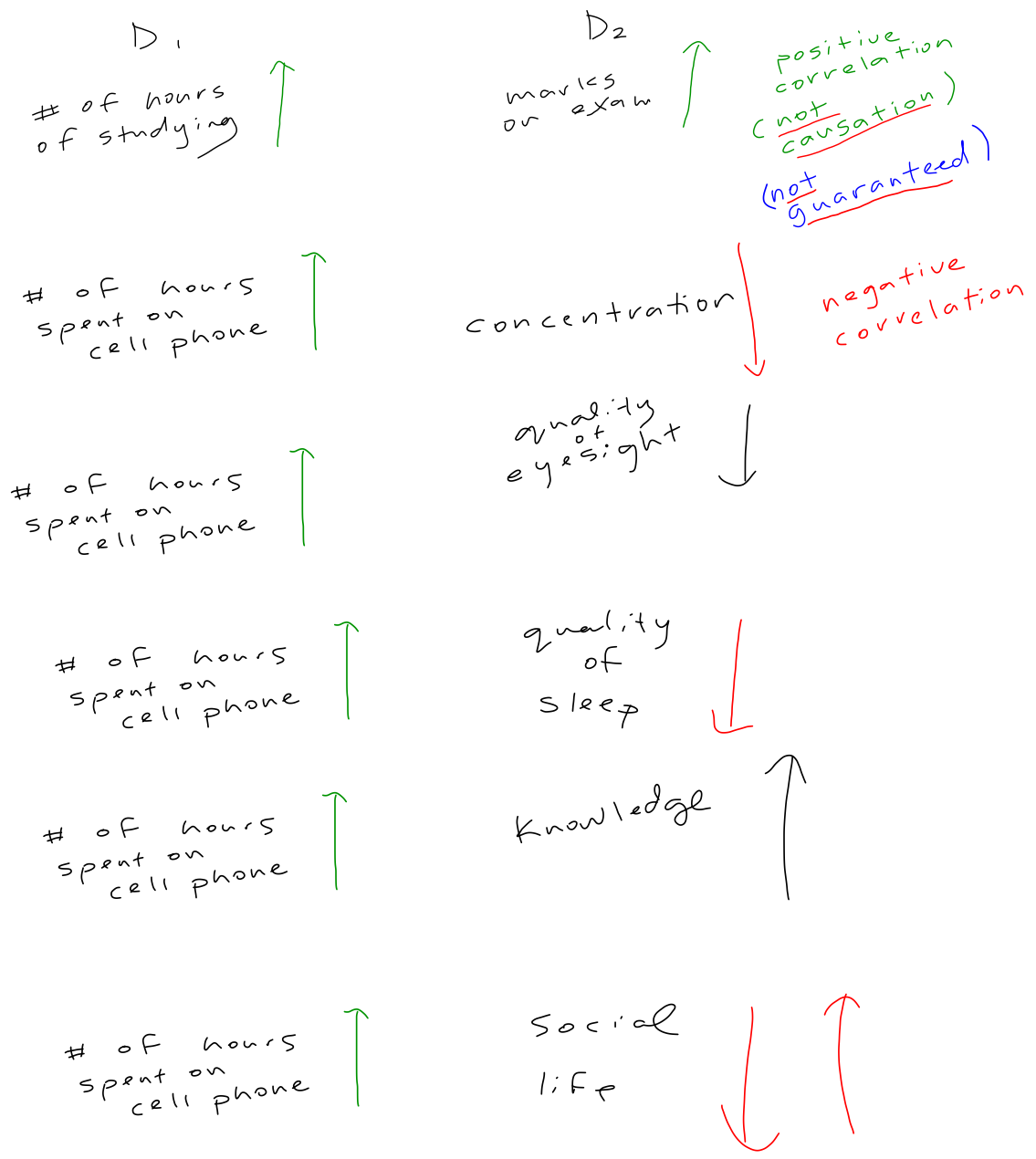


Lesson 4: Calculating the Correlation Coefficient (r) Between 2 Distributions

Definition: a correlation is a link/connection between two separate distributions



2 distributions. Are they linked?

Tennis training

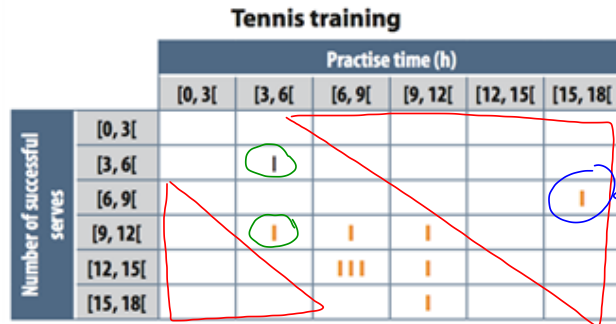
Player	Practise time (h)	Number of successful serves
Alex	4	5
Benoit	5	10
Cloé	6	12
Dylan	6	14
Éric	7	9
France	8	13
Gaël	9	11
Henri	9	15
Isaak	11	14
Julie	15	6

There are 2  
ways to visualize  
...

2<sup>st</sup> A way to visualize if there's a correlation: constructing a contingency table

**Tennis training**

Player	Practise time (h)	Number of successful serves
Alex	4	5
Benoit	5	10
Cl�e	6	12
Dylan	6	14
�ric	7	9
France	8	13
Ga�l	9	11
Henri	9	15
Isaak	11	14
Julie	15	6



if corners are empty there's a correlation.  
 - outlier - eliminate it.

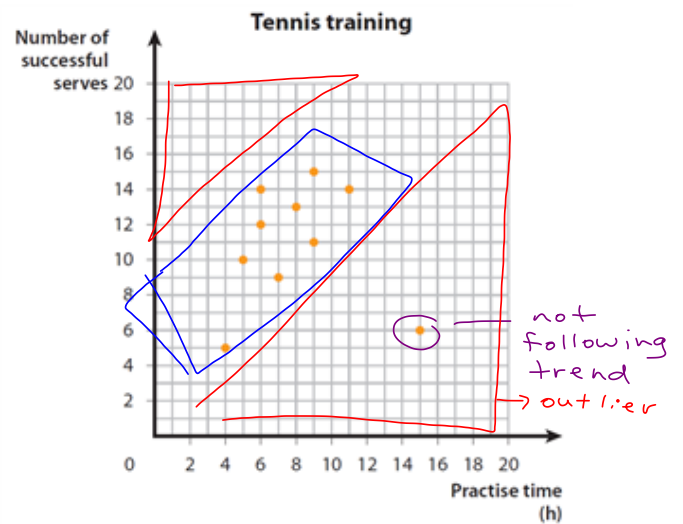
Tennis training

Player	Practise time (h)	Number of successful serves
Alex	4	5
Benoît	5	10
Cloé	6	12
Dylan	6	14
Éric	7	9
France	8	13
Gaël	9	11
Henri	9	15
Isaak	11	14
Julie	15	6

2<sup>nd</sup> way to visualize if there's a correlation: create a scatter plot

Tennis training

Player	Practise time (h)	Number of successful serves
Alex	4	5
Benoît	5	10
Cloé	6	12
Dylan	6	14
Éric	7	9
France	8	13
Gaël	9	11
Henri	9	15
Isaak	11	14
Julie	15	6



Use calculator to calculate strength of correlation

Player	Practise time (h)	Number of successful serves
Alex	4	5
Benoit	5	10
Cloé	6	12
Dylan	6	14
Éric	7	9
France	8	13
Gaël	9	11
Henri	9	15
Isaak	11	14
Julie	15	6

Find  $r$

Correlation coefficient.

(google instructions specific to your calculator model)